



and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1807	100.0	340	2	AAW68520	Aaw68520 Salmonid-
2	1807	100.0	340	2	AAY17440	Aay17440 SB transp
3	1807	100.0	340	4	AAY72024	Aay72024 SB (Sleep
4	1807	100.0	340	5	AAG80595	Aag80595 Sleeping
5	1807	100.0	340	5	ABB07570	Abb07570 SB transp
6	1807	100.0	340	6	ABB99719	Abb99719 Amino aci
7	1807	100.0	340	7	ADD41924	Add41924 Amino aci
8	1807	100.0	340	10	ADW22678	Adw22678 Sleeping
9	1807	100.0	340	10	ADW03280	Adw03280 Sleeping
10	1807	100.0	340	11	AEF81669	Aef81669 Salmonid
11	1807	100.0	340	12	AEP47151	Aep47151 SB transp
12	1802	99.7	340	7	ADK17480	Adk17480 Salmonid
13	1787	98.9	340	11	AEF81683	Aef81683 Salmonid
14	1786	98.8	340	4	AAB47211	Aab47211 Sleeping
15	1786	98.8	340	10	ADW46561	Adw46561 Salmonid
16	1786	98.8	340	12	AEL92393	Ael92393 Sleeping
17	1784	98.7	340	11	AEF81684	Aef81684 Salmonid
18	1783	98.7	340	10	ADW46599	Adw46599 Salmonid
19	1783	98.7	340	10	ADW46649	Adw46649 Salmonid
20	1783	98.7	340	10	ADW46612	Adw46612 Salmonid
21	1783	98.7	340	10	ADW46621	Adw46621 Salmonid
22	1783	98.7	340	10	ADW46613	Adw46613 Salmonid
23	1782	98.6	340	10	ADW46626	Adw46626 Salmonid
24	1782	98.6	340	10	ADW46644	Adw46644 Salmonid
25	1782	98.6	340	10	ADW46614	Adw46614 Salmonid
26	1782	98.6	340	10	ADW46597	Adw46597 Salmonid
27	1782	98.6	340	10	ADW46587	Adw46587 Salmonid
28	1781	98.6	340	10	ADW46656	Adw46656 Salmonid
29	1781	98.6	340	10	ADW46591	Adw46591 Salmonid
30	1781	98.6	340	10	ADW46643	Adw46643 Salmonid
31	1781	98.6	340	10	ADW46658	Adw46658 Salmonid
32	1781	98.6	340	10	ADW46588	Adw46588 Salmonid
33	1781	98.6	340	10	ADW46605	Adw46605 Salmonid
34	1781	98.6	340	10	ADW46603	Adw46603 Salmonid
35	1781	98.6	340	10	ADW46653	Adw46653 Salmonid
36	1781	98.6	340	10	ADW46607	Adw46607 Salmonid
37	1781	98.6	340	10	ADW46615	Adw46615 Salmonid
38	1781	98.6	340	10	ADW46617	Adw46617 Salmonid
39	1781	98.6	340	10	ADW46590	Adw46590 Salmonid
40	1781	98.6	340	10	ADW46593	Adw46593 Salmonid
41	1781	98.6	340	10	ADW46608	Adw46608 Salmonid
42	1781	98.6	340	10	ADW46645	Adw46645 Salmonid
43	1781	98.6	340	10	ADW46601	Adw46601 Salmonid
44	1781	98.6	340	10	ADW46610	Adw46610 Salmonid
45	1781	98.6	340	10	ADW46630	Adw46630 Salmonid

## ALIGNMENTS

## RESULT 1

AAW68520

ID AAW68520 standard; protein; 340 AA.

XX

AC AAW68520;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-JAN-1999 (first entry)  
 XX  
 DE Salmonid-type Tc1-like transposase protein.  
 XX  
 KW Transposase; salmonid; Tc1-like family; homology; vertebrate; transposon;  
 KW zebrafish; gene transfer; gene therapy; inverted repeat; direct repeat;  
 KW genome.  
 XX  
 OS transposons.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 11. .61  
 FT /note= "Paired-like domain with Leucine Zipper"  
 FT Domain 104. .105  
 FT /note= "form part of a bipart nuclear localisation  
 FT signal"  
 FT Domain 116. .120  
 FT /note= "form part of a bipart nuclear localisation  
 FT signal"  
 FT Domain 151. .161  
 FT /note= "DD(34)E box"  
 FT Region 183. .198  
 FT /note= "Glycine-rich box"  
 FT Domain 241. .251  
 FT /note= "DD(34)E box"  
 FT Domain 268. .282  
 FT /note= "DD(34)E box"  
 XX  
 PN WO9840510-A1.  
 XX  
 PD 17-SEP-1998.  
 XX  
 PF 11-MAR-1998; 98WO-US004687.  
 XX  
 PR 11-MAR-1997; 97US-0040664P.  
 PR 28-JUL-1997; 97US-0053868P.  
 PR 13-NOV-1997; 97US-0065303P.  
 XX  
 PA (MINU ) UNIV MINNESOTA.  
 XX  
 PI Hackett PB, Ivics Z, Izsvak Z, Caldovic L;  
 XX  
 DR WPI; 1998-531525/45.  
 DR N-PSDB; AAV33841.  
 XX  
 PT New isolated transposase gene - used to develop products for the  
 PT integration of nucleic acid fragments into the nucleic acid of cells,  
 PT particularly for gene therapy.  
 XX  
 PS Disclosure; Fig 2B; 79pp; English.  
 XX  
 CC This sequence represents a novel transposase protein of the salmonid-type  
 CC Tc1-like family. The coding sequence was reconstructed based on sequence  
 CC homology between members of the Tc1-like family of vertebrate  
 CC transposons, especially based on the sequence of a number of fish, e.g.  
 CC zebrafish transposons. The transposase can be used in a gene transfer  
 CC method, especially for gene therapy treatment, by flanking a heterologous  
 CC gene to be expressed in a cell with the inverted or direct repeat

CC sequences to which the transposase binds. This allows the transposase-  
CC mediated transposition of the heterologous into the genome of the host  
CC cell. (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 340 AA;

Query Match 100.0%; Score 1807; DB 2; Length 340;  
Best Local Similarity 100.0%; Pred. No. 2.4e-156;  
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGKSKEISQDLRKKIVDLHKSGSSLGAISKRLKVPRSSVQTIVRKYKHHGTTQPSYRSGR	60
Db	1	MGKSKEISQDLRKKIVDLHKSGSSLGAISKRLKVPRSSVQTIVRKYKHHGTTQPSYRSGR	60
Qy	61	RRVLSPRDERTLVRKVQINPRTTAKDLVKMLEETGTKVSISTVKRVLYRHNLKGRSARKK	120
Db	61	RRVLSPRDERTLVRKVQINPRTTAKDLVKMLEETGTKVSISTVKRVLYRHNLKGRSARKK	120
Qy	121	PLLQNRHKKARLRFATAHGDKDRTFWRNVLWSDETKIELFGHNDHRYVWRKKGEACKPKN	180
Db	121	PLLQNRHKKARLRFATAHGDKDRTFWRNVLWSDETKIELFGHNDHRYVWRKKGEACKPKN	180
Qy	181	TIPTVKHGGGSIMLWGCFAAGGTGALHKIDGIMRKENYVDILKQHLKTSVRKLKLGRKWV	240
Db	181	TIPTVKHGGGSIMLWGCFAAGGTGALHKIDGIMRKENYVDILKQHLKTSVRKLKLGRKWV	240
Qy	241	FQMDNDPKHTSKVVAKWLKDNKVKVLEWPSQSPDLNPIENLWAEKKRVRARRPTNLTQL	300
Db	241	FQMDNDPKHTSKVVAKWLKDNKVKVLEWPSQSPDLNPIENLWAEKKRVRARRPTNLTQL	300
Qy	301	HQLCQEEWAKIHPTYCGKLVEGYPKRLTQVKQFKGNATKY	340
Db	301	HQLCQEEWAKIHPTYCGKLVEGYPKRLTQVKQFKGNATKY	340